APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR		ATTORNEY DOCKET NO.
08/825,534	03/28/97	YOUNG	J	06998/022001

LM91/0422

FISH AND RICHARDSON
601 THIRTEENTH STREET NW
WASHINGTON DC 20005

EXAMINER
LERNER, M

ART UNIT	PAPER NUMBER
2741	14

DATE MAILED: 04/22/99

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner of Patents and Trademarks

Office Action Summary

☒ Responsive to communication(s) filed on *Mar 10, 1999*

☒ This action is **FINAL**.

☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11; 453 O.G. 213.

A shortened statutory period for response to this action is set to expire ___*three*___ month(s), or thirty days, whichever is longer, from the mailing date of this communication. Failure to respond within the period for response will cause the application to become abandoned. (35 U.S.C. § 133). Extensions of time may be obtained under the provisions of 37 CFR 1.136(a).

Disposition of Claims

☒ Claim(s) *1-6, 8-25, and 27-31* is/are pending in the application.

    Of the above, claim(s) _____ is/are withdrawn from consideration.

☐ Claim(s) _____ is/are allowed.

☒ Claim(s) *1-6, 8-25, and 27-31* is/are rejected.

☐ Claim(s) _____ is/are objected to.

☐ Claims _____ are subject to restriction or election requirement.

Application Papers

☐ See the attached Notice of Draftsperson's Patent Drawing Review, PTO-948.

☐ The drawing(s) filed on _____ is/are objected to by the Examiner.

☐ The proposed drawing correction, filed on _____ is ☐approved ☐disapproved.

☐ The specification is objected to by the Examiner.

☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. § 119

☐ Acknowledgement is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d).

    ☐ All ☐ Some* ☐None   of the CERTIFIED copies of the priority documents have been

        ☐ received.

        ☐ received in Application No. (Series Code/Serial Number) _____ .

        ☐ received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

    *Certified copies not received: _____ .

☐ Acknowledgement is made of a claim for domestic priority under 35 U.S.C. § 119(e).

Attachment(s)

☐ Notice of References Cited, PTO-892

☐ Information Disclosure Statement(s), PTO-1449, Paper No(s). _____

☐ Interview Summary, PTO-413

☐ Notice of Draftsperson's Patent Drawing Review, PTO-948

☐ Notice of Informal Patent Application, PTO-152

--- *SEE OFFICE ACTION ON THE FOLLOWING PAGES* ---

DETAILED ACTION

Claim Rejections - 35 USC § 102

1.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the

basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless --
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

2.   Claims 1 to 6 and 12 are rejected under 35 U.S.C. 102(b) as being anticipated by *Gould et*

al. ('707).

Regarding independent claim 1, *Gould et al. ('707)* discloses:

"performing speech recognition on an utterance to produce a recognition result for the

utterance" -- DragonDictate receives an utterance whose best scoring word is a Choice Command

(column 12, lines 48 to 54);

"identifying a correction command in the recognition result for the utterance" -- a

correction command corresponding to an utterance "Choose-N," "Scratch-that" is recognized

(column 12, lines 55 to 60);

"identifying corrected text from a portion of the recognition result for the utterance" -- an

number of backspace keystrokes are erased to correct the text (column 12, lines 60 to 68);

"wherein the correction command indicates that the portion of the recognition results

comprises a pronunciation of a word to be corrected" -- if a user then says an utterance other than

a command, i.e. an utterance for training, the systems enters a Confirmed Training Only Routine

where a new entry in the Oops Buffer is used to update word and language models (column 14,

lines 14 to 50).

Regarding claim 2, *Gould et al. ('707)* discloses "replacing previously-generated incorrect

text with corrected text" -- the most recent utterance in the Oops Buffer is corrected by means of

the Choice Commands (column 13, lines 11 to 15).

Regarding claim 3, *Gould et al. ('707)* discloses "wherein the step of identifying

corrected text includes searching a dictionary using the portion of the recognition results" -- the

vocabulary file (".VOC file") and the user file (".USR file") are implicitly used to as dictionaries

representing phonetic spellings to identify an utterance (column 10, lines 17 to 24 and

column 10, line 64 to column 11, line 2).

Regarding claim 4, *Gould et al. ('707)* discloses "wherein the step of identifying

corrected text comprises identifying corrected text from a portion of the recognition result for the

utterance and from a recognition result for a second utterance" -- the "Left-1" and "Right-1"

commands move the word in the Oops Buffer left or right by one word so that this word may be

corrected (column 14, lines 5 to 9).

Regarding claim 5, *Gould et al. ('707)* discloses "wherein the second utterance precedes

the utterance" -- the "Left-1" command moves the Oops buffer to a preceding utterance.

Regarding claim 6, *Gould et al. ('707)* discloses "wherein the second utterance follows

the utterance" -- the "Right-1" command moves the Oops buffer to a following utterance.

Regarding claim 12, *Gould et al. ('707)* discloses "automatically selecting the previously-generated incorrect text to be replaced" -- text to be corrected is automatically highlighted (Figures 36 to 63).

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 8 to 11 and 13 to 24 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Gould et al. ('707)* in view of *Roberts et al.*

Concerning claim 8, *Gould et al. ('707)* does not disclose identifying corrected text using "confused pronunciation matching." However, *Roberts et al.* uses a phonetic dictionary 500a, where each of the word entries is associated with a phonetic spelling using acoustic node models (column 18, lines 43 to column 19, line 13 and Figure 8) for the purpose of improving performance of speech recognition by taking into account preceding and following phonemes. The phonetic dictionary 500a of *Roberts et al.* is used with correction commands (column 21, lines 41 to 56) and performs "confused pronunciation matching." *Gould et al. ('707)* and *Roberts et al.* belong to the same field of endeavor. It would have been obvious to one of ordinary skill in the art to use a phonetic dictionary to perform "confused pronunciation

matching" as taught by *Roberts et al.* for the purpose of improving recognition results in a correction mode.

Concerning claim 9, the phonetic dictionary 500a of *Roberts et al.* is "confused pronunciation dictionary."

Concerning claim 10, the phonetic dictionary 500a of *Roberts et al.* is used to search for confused pronunciation matches.

Concerning claim 11, the phonetic dictionary 500a of *Roberts et al.* is constructed with acoustic node models, or as a "phonetic tree" (column 18, lines 60 to 65 and Figure 8).

Concerning claim 13, *Roberts et al.* teaches "re-recognition" of corrected text during the correction process (step 272).

Concerning claim 14, *Roberts et al.* generates a list of words corresponding to the entries in phonetic dictionary 500a for text to be corrected in correction mode (column 21, lines 41 to 56 and Figures 10 to 24).

Concerning claim 15, *Roberts et al.* teaches "re-recognition" of corrected text during the correction process (step 272) from a restricted phonetic vocabulary (column 21, lines 53 to 56).

Concerning claim 16, *Roberts et al.* displays a list of words corresponding to the entries in phonetic dictionary 500a for a user to select with a correction command (column 21, lines 41 to 56 and Figures 10 to 24).

Concerning claim 17, *Roberts et al.* discloses spelling commands "starts alpha," "starts beta," etc. indicating a portion of a recognition result to be corrected (column 19, line 57 to column 20, line 19).

Concerning claim 18, dictionary 500 of *Roberts et al.* consists of an alphabetical listing of word spellings (column 18, lines 46 to 51). Dictionary 500 is used to perform "confused spelling matching" in a correction mode (column 20, lines 52 to 62).

Concerning claim 19, backup dictionary 500 of *Roberts et al.* is a "confused spelling dictionary" (column 20, lines 52 to 62).

Concerning claim 20, backup dictionary 500 of *Roberts et al.* is a "confused spelling dictionary" that is searched during correction (column 20, lines 52 to 62).

Concerning claim 21, *Roberts et al.* generates a list of words corresponding to the entries in spelling dictionary 500 for a user to select with a correction command (Figures 10 to 24).

Concerning claim 22, *Roberts et al.* teaches "re-recognition" of corrected text during the correction process from backup dictionary 500 (column 21, lines 5 to 16).

Concerning claim 23, *Roberts et al.* displays a list of words corresponding to the entries in spelling dictionary 500 for a user to select with a correction command (Figures 10 to 24).

Concerning claim 24, *Roberts et al.* discloses:

"using an active vocabulary when performing speech recognition" -- TEXTMODE and EDITMODE use different active vocabularies (column 8, lines 51 to 54);

"using a backup dictionary when identifying the corrected text" -- a backup dictionary is used in EDITMODE (column 20, lines 52 to 62);

"if the active vocabulary does not contain the corrected text, adding the corrected text to the active vocabulary" -- a new word is added to the vocabulary through a Definition Window (column 20, line 63 to column 21, line 4).

5.   Claims 25 and 27 to 31 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Roberts et al.* in view of *Junqua.*

Concerning claim 25, *Roberts et al.* discloses:

"performing speech recognition on an utterance to produce recognition results" -- a dictation program detects speech inputted by a user corresponding to a letter command (column 19, lines 46 to 56);

"identifying a spelling command in the recognition results, wherein the spelling command indicates that a portion of the utterance comprises a spelling" -- spelling commands "starts alpha," "starts beta," etc., of portions of an utterance are identified (column 19, line 57 to column 20, line 6); and

"producing the spelling by searching a dictionary using the recognition results" -- spellings are searched through a limited vocabulary dictionary (column 20, lines 44 to 51).

Roberts et al. suggests that spelling may be confusingly similar, i.e. "confused spelling matching" (column 20, line 18), but does not expressly disclose "commonly-confused letters are

treated as a single letter to identify the spelling corresponding to the portion of the utterance."

However, *Junqua* teaches confused spelling based upon how confusable particular letters are

with respect to one another, e.g. m and n, p and t, or b and d. See column 6, lines 26 to 67. It

would have been obvious to one of ordinary skill in the art to use state tying of confusable letters

as taught by *Junqua* for the purpose of improving recognition accuracy by pruning the number of

paths during a beam search.

Concerning claim 27, dictionary 500 of *Roberts et al.* consists of an alphabetical listing of

word spellings, or "confused spelling dictionary" (column 18, lines 46 to 51). Dictionary 500 is

used to perform "confused spelling matching" in a correction mode (column 20, lines 52 to 62).

Concerning claim 28, dictionary 500 of *Roberts et al.* consists of an alphabetical listing of

word spellings (column 18, lines 46 to 51). Dictionary 500 is used to perform "confused spelling

matching" in a correction mode (column 20, lines 52 to 62).

Concerning claim 29, *Roberts et al.* generates a list of words corresponding to the entries

in spelling dictionary 500 for a user to select with a correction command (Figures 10 to 24).

Concerning claim 30, *Roberts et al.* displays a list of words corresponding to the entries

in spelling dictionary 500 for a user to select with a correction command (Figures 10 to 24).

Concerning claim 31, *Junqua* discloses that confused spelling matching uses vowel

sounds, fricatives, affricatives, plosives and nasals ("characteristics of a speaker's

pronunciation") to provide distinguishing features between confusable letters.

Response to Arguments

6.   Applicants' arguments filed March 10, 1999 have been fully considered but they are not

persuasive.

Regarding the rejection of claims 1 to 6 and 12 as being anticipated by *Gould et al.*

('707), Applicants state that the "CHOOSE-N" command could arguably be said to identify

corrected text (Remarks, Page 2), but that the delete commands do not identify corrected text

(Remarks, Page 3). It is agreed, insofar as the limitation in the third clause of claim 1

("identifying corrected text from a portion of the recognition result for the utterance"), that the

"CHOOSE-N" command automatically identifies corrected text after the choice command has

been identified. This is illustrated by *Gould et al. ('707)* in Figures 46 to 50 where

DragonDictate misrecognizes "vary" for "very". Once the user says "choose 3" (Figure 49),

DragonDictate uses the third choice in choice window 664 to correct the last word recognized

("vary"). Because a "SCRATCH THAT"command identifies and deletes incorrect text, i.e. the

last text word recognized, it may arguably be maintained that identification and removal of

incorrect text corrects the text in some cases ("identifying corrected text"). However, the

"CHOOSE-N" command is a better illustration of this limitation.

Next, Applicants maintain that *Gould et al. ('707)* does not meet the limitation that the

correction command "indicates that the portion of the recognition results comprises a

pronunciation of a word to be corrected." At the outset, Applicants' comments with respect to

"an utterance" and "a recognition result for the utterance" assume too restrictive an interpretation

of these terms. Broadly speaking, "an utterance" may be construed to be continuous dictation of

original text as well as correction commands with associated corrected text. "Recognition

results" are seamlessly produced for both dictation of original text and correction commands in

the dictation system of *Gould et al. ('707)*. In the example of Figure 49, a user first says "very,"

and then sometime later says "choose 3." The utterance therefore comprises "very . . . choose 3."

Gould et al. ('707) performs speech recognition on the utterance to produce a recognition result

for the utterance "very . . . choose 3." For the utterance "very," DragonDictate produces the

(mis)recognition result "vary." For the utterance "choose 3," a CHOOSE-N correction command

is identified, and the program automatically corrects the last recognition result ("vary") with

corrected text from the third item in correction window 664.

The pronunciation of a word to be corrected is utilized by an adaptive training subroutine

for the recognition results of "vary," the word to be corrected. Adaptive training is utilized

automatically to ensure that the program will not continue to misrecognize a word in the future

(Figure 5, step 392). *Gould et al. ('707)* calls the speech information signal for a command or

dictated word a "token." A "token" is the sound of a word, as distinguished from the word as

text. A "token" is therefore a "pronunciation of a word." During dictation, the last token

recognized is automatically stored for the entry in the OOPS buffer to correct the word if the

word turns out to have been misrecognized (Figure 5, step 400). The token corresponding to the

last entry in the OOPS buffer is thus "a pronunciation of a word to be corrected." Unless a user

specifies from the beginning by setting a Confirmed Training Only Flag to true (Figure 6, step

468), DragonDictate automatically uses a token in the OOPS buffer for adaptive training in the

course of dictation to continually update the language models (Figure 5, steps 394 to 398 and

Figure 12). However, when a CHOOSE-N command is selected, the program sets the Confirmed

Training Flag Only to false (Figure 5, step 254) so that adaptive training is always performed

with the token of the entry in the OOPS buffer for this command (Figure 5, step 256);

DragonDictate thus ensures that adaptive training is performed when a word has been

misrecognized. The correction command CHOOSE-N indicates that the language model for the

confirmed utterance (the corrected word "very") and the first word choice (the incorrect word

"vary") are updated with the token in OOPS buffer (Figure 5, step 256). Hence, the correction

command CHOOSE-N indicates that a portion of the recognition results includes a pronunciation

of a word to be corrected, i.e. the token for "very." Applicants' position that *Gould et al. ('707)*

omits disclosing a correction command which indicates that a portion of the recognition results

comprises a pronunciation of a word to be corrected is thereby traversed.

Regarding the rejection of claims 25 and 27 to 30 as being obvious over *Roberts et al.* in

view of *Junqua*, Applicants state that there is no motivation to combine these references. In

traversing the rejection, Applicants maintain that *Roberts et al.* would have no use for confused

spelling matching because this reference includes letter commands, "starts_comletter" (e.g.

"starts_alpha", "starts_beta"). Similarly, Applicants state that *Junqua* teaches recognition of

spoken letters is improved with the use of a phonetic alphabet (A-Alpha, B-Baker, C-Charlie).

The point of the teaching of *Junqua*, however, is that confused spelling matching makes

letter commands or a phonetic alphabet unnecessary. Assume a word "invention" has been

misrecognized as "inversion." With the letter commands of *Roberts et al.*, a user would have to

speak the command "starts_eye". This produces a much longer list than if the user speaks a

command "starts invent". See Figure 15 of *Roberts et al.* A longer list of potential candidates

produces slower correction. *Junqua* suggests that confused spelling matching decreases the

response time by producing a shorter list of candidates. See column 1, lines 50 to 67 of *Junqua*.

Thus, the motivation taught by *Junqua* is to increase speed while maintaining accuracy.

Conclusion

7.   **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time

policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within TWO

MONTHS of the mailing date of this final action and the advisory action is not mailed until after

the end of the THREE-MONTH shortened statutory period, then the shortened statutory period

will expire on the date the advisory action is mailed, and any extension fee pursuant to 37

CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event,

however, will the statutory period for reply expire later than SIX MONTHS from the mailing

date of this final action.

8.   Any inquiry concerning this communication or earlier communications from the examiner

should be directed to *Martin Lerner* whose telephone number is (703) *308-9064*.

The fax phone number for the organization where this application or proceeding is

assigned is (703) 305-9508.

Any inquiry of a general nature or relating to the status of this application or proceeding

should be directed to the receptionist whose telephone number is (703) 305-4800.

DAVID R. HUDSPETH
SUPERVISORY PATENT EXAMINER
GROUP 2700

ml

April 16, 1999